```
                          GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:            July 4, 2001, 18:30:48 ; Search time 3525.91 Seconds
                                            (without alignments)
                                            9884.731 Million cell updates/sec

Title:
Perfect score:     3687
Sequence:          1 atggatggaaaaagctgt.........ttacattctatattaagtaa 3687

Scoring table:     IDENTITY_NUC
                   Gapop 10.0 , Gapext 1.0

Searched:          10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters:      20456230

Post-processing:   Minimum Match 0%
Maximum DB seq length: 2000000000       Maximum Match 100%
Minimum DB seq length: 0                Listing first 45 summaries

Database :
```

```
EST: *
  1: gb_est1: *          44: em_esthum10: *
  2: gb_est2: *          45: em_esthum11: *
  3: gb_est3: *          46: em_esthum12: *
  4: gb_est4: *          47: em_esthum13: *
  5: gb_est5: *          48: em_esthum14: *
  6: gb_est6: *          49: em_esthum15: *
  7: gb_est7: *          50: em_esthum16: *
  8: gb_est8: *          51: em_esthum17: *
  9: gb_est9: *          52: em_esthum18: *
 10: gb_est10: *         53: em_esthum19: *
 11: gb_est11: *         54: em_esthum20: *
 12: gb_est12: *         55: em_esthum21: *
 13: gb_est13: *         56: em_esthum22: *
 14: gb_est14: *         57: em_esthum23: *
 15: gb_est15: *         58: em_esthum24: *
 16: gb_est16: *         59: em_esthum25: *
 17: gb_est17: *         60: em_esthum26: *
 18: gb_est18: *         61: em_esthum27: *
 19: gb_est19: *         62: em_esthum28: *
 20: gb_est20: *         63: em_estin1: *
 21: gb_est21: *         64: em_estin2: *
 22: gb_est22: *         65: em_estin3: *
 23: gb_est23: *         66: em_estin4: *
 24: gb_est24: *         67: em_estin5: *
 25: gb_est33: *         68: em_estom1: *
 26: gb_est34: *         69: em_estom2: *
 27: gb_est35: *         70: em_estov1: *
 28: gb_est36: *         71: em_estov2: *
 29: gb_est37: *         72: em_estpl1: *
 30: gb_est38: *         73: em_estpl2: *
 31: gb_est39: *         74: em_estpl13: *
 32: gb_est40: *         75: em_estpl14: *
 33: em_estba: *         76: em_estpl15: *
 34: em_estfun: *        77: em_estpl16: *
 35: em_esthum1: *       78: em_estpl17: *
 36: em_esthum2: *       79: em_estpl18: *
 37: em_esthum3: *       80: em_estpl19: *
 38: em_esthum4: *       81: em_estpl110: *
 39: em_esthum5: *       82: em_estpl11: *
 40: em_esthum6: *       83: em_estro1: *
 41: em_esthum7: *       84: em_estro2: *
 42: em_esthum8: *       85: em_estro3: *
 43: em_esthum9: *       86: em_estro4: *
                         87: em_estro5: *
                         88: em_estro6: *
                         89: em_estro7: *
                         90: em_estro8: *
                         91: em_estro9: *
                         92: em_estro10: *
                         93: em_estro11: *
                         94: em_estro12: *
                         95: em_estro13: *
                         96: em_estro14: *
                         97: em_estro15: *
                         98: em_estro16: *
                         99: em_estro17: *
                        100: em_estro18: *
                        101: em_estro19: *
                        102: em_estro20: *
                        103: gb_est25: *
                        104: gb_est26: *
                        105: gb_est27: *
                        106: gb_est28: *
                        107: gb_est29: *
                        108: gb_est30: *
                        109: gb_est31: *
                        110: gb_est32: *
                        111: gb_est41: *
                        112: gb_est42: *
                        113: gb_est43: *
                        114: gb_est44: *
                        115: gb_est45: *
                        116: gb_est46: *
                             gb_est47: *
```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est75:*
137: gb_est76:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est67:*
161: gb_est68:*
162: gb_est69:*
163: gb_est70:*
164: gb_est71:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est91:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthum29:*
177: em_esthum30:*
178: em_esthum31:*
179: em_esthum32:*
180: em_esthum3:*
181: em_estom3:*
182: em_estpl11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*

190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun1:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: gb_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

RESULT 1
A2548467
LOCUS   AZ548467   908 bp   DNA   GSS   14-NOV-2000
DEFINITION ENTEK30TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ548467
VERSION   AZ548467.1 GI:11172102
KEYWORDS   GSS.
SOURCE   Entamoeba histolytica.
ORGANISM   Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 908)
AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
TITLE   Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
JOURNAL   Unpublished (2000)
COMMENT   Contact: Brendan J Loftus

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: bjloftus@tigr.org

FEATURES   Location/Qualifiers
  source   1..908
             /organism="Entamoeba histolytica"
             /strain="HM1:IMSS"
             /db_xref="taxon:5759"
             /clone_lib="Entamoeba histolytica Sheared DNA"
             /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
             Institute for Genomic Research (TIGR), Rockville, MD.
             Genomic DNA isolated from broth cultures of E. histolytica
             using a method described by Clark and Diamond (Clark,
             C.G., and Diamond, L.S. (1993) Entamoeba histolytica; a
             method for isolate identification. Exp. Parasitol.
             77:450.). The DNA was mechanically sheared to give a
             tight size distribution (~2 kb). The v + 1 method used for
             the library construction is described in detail in Smith,
             H.O. and Venter, J.C. (Making small insert libraries for
             whole genome shotgun sequencing projects. In Genome
             Sequencing: A Practical Approach, eds. M. Vaudin and B.
             Barell, Oxford University Press, 1999)."
BASE COUNT   434 a   16 c   289 g   169 t
ORIGIN

SUMMARIES

	Result No.	Score	Query Match	Length	DB ID	Description
	1	72.2	2.0	908	245	AZ548467 ENTEK30TR
	2	68.2	1.8	877	245	AZ531291 ENTBQ34TR
	3	68.2	1.8	912	245	AZ551092 ENTFJ22TF
	4	66.8	1.8	816	245	AZ535744 ENTCO25TR
c	5	66	1.8	843	245	AZ551618 ENTDV54TR
c	6	63.6	1.7	880	245	AZ529191 ENTEV68TR
c	7	62	1.7	849	245	AZ546009 ENTFW53TF
c	8	61	1.7	891	248	AZ683582 ENTKK47TR
c	9	60	1.6	900	245	AZ549980 ENTDD94TF
c	10	59.8	1.6	997	219	CN501344P Drosophil
	11	58.8	1.6	494	258	FR048073 Fugu rubr
	12	58.8	1.6	905	245	AZ550256 ENTEV58TR
	13	58.4	1.6	912	245	AZ527724 ENTBD71TR
c	14	57.4	1.6	938	247	AZ676392 ENTHY47TR
c	15	57.4	1.6	879	245	AZ550718 ENTEM36TF
	16	56.8	1.5	476	165	BE224646 kp59a06.y
	17	56.4	1.5	910	245	AZ544844 ENTEV52TF
	18	55.8	1.5	1159	219	CN5015XR Drosophil
	19	55.8	1.5	773	220	CN501VTG Drosophil
	20	55.4	1.5	735	221	CN504NSM Tetraodon
	21	55.2	1.5	917	219	CN017SL Drosophil
	22	54.2	1.5	1076	222	CN505HXN Tetraodon
	23	54.2	1.5	1282	80	BF264918 HV_CEa001
	24	53.8	1.5	619	258	FR047601 Fugu rubr
	25	53.8	1.5	906	245	AZ547385 ENTGK54TR
	26	53.8	1.5	918	247	AZ677162 ENTIU66TF
	27	53.4	1.5	560	219	CN50C13 Drosophil
c	28	53.4	1.4	1101	219	CN501807 Drosophil
c	29	53	1.4	1025	219	CN014J2 Drosophil
c	30	53	1.4	1147	256	B13042 T30M24-Sp6.
c	31	52.4	1.4	501	258	FR048173 Fugu rubr
c	32	51.8	1.4	959	219	CN500655 Drosophil
c	33	51.8	1.4	969	219	CN501DL Drosophil
c	34	51.8	1.4	1101	219	CN500HX9 Drosophil
c	35	51.6	1.4	438	174	BG226596 Drosophil
c	36	51.6	1.4	470	222	FR018463 Drosophil
c	37	51	1.4	911	247	AZ668666 ENTIM95TR
c	38	51	1.4	1042	219	CN50148K Drosophil
c	39	51	1.4	1184	221	CN504P4P Drosophil
c	40	50.8	1.4	421	137	BE580431 Tetraodon
c	41	50.8	1.4	460	174	BG224667 Drosophil
c	42	50.8	1.4	888	245	AZ549065 ENTGE24TR
c	43	50.8	1.4	1101	219	CN500039G Drosophil
	44	50.6	1.4	539	114	AW329517 N200762e
	45	50.6	1.4	1101	219	CN500EVL Drosophil

ALIGNMENTS

Query Match 2.0%; Score 72.2; DB 245; Length 908;
Best Local Similarity 44.6%; Pred. No. 2.8e-06;
Matches 333; Conservative 0; Mismatches 408; Indels 6; Gaps 1;

```
Qy  195  taacgatgtatatgctgaatacaacaaagcgaaaaaacgataccgtgatgcggtagcatt  254
         | ||| | || ||||||| | |||  | | ||||||||||||||||||||||||||||||
Db   42  TGATGATGAAGAAGATGATGAAGATGATGAAGAAGATGATGAAGATGATGAAGAAGATGAA 101

Qy  255  agtgaataaagcaggtgggcgcgaaaaagaagacgcttacttagctgattacaaaagaata  314
         | ||| | || ||  |   | | |  ||| ||||||||||||||||||||||||||
Db  102  AGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGATGATGAAGACGA  161

Qy  315  tgaaacttacgtttcaaagcaaaccctaaatctggcgaagctcgtgtagcaacttacat  374
         |||| | ||  ||| |||  |||  ||||| || | || |||| |||| | |||||||
Db  162  TGAAGACGACGAAATATGAATTAGAAGATGATGAAGATGATGAAGATGATGAAGATGATGA  221

Qy  375  cgatgcttacaactatgcaacaaaattagaacgaaatgcgccaagagagactagagcgtgt  434
         ||| | |||  | | || |  |||||||| |  ||||  ||||||||| |||||| | |||
Db  222  AGATGATGAAGACGATGAAGAAGACGAATATGAATT------AGAAGATGATGAAGATGA  275

Qy  435  tcaagcaaaagatttagaaaaagcagaacaatactacacaaaatccttatgaattaa  494
         ||||| ||||||| ||| |||||||| | |||| || || |||| ||||||||||
Db  276  TGAAGAAGATGATGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGAAGA  335

Qy  495  aactcgcacagtcatttagtcgtatatggtaaacaacaacctgtgattacttcgctc  554
         | || |  || ||| | ||| | |||| | |||||| ||| || |||| |||||||
Db  336  AGATGAAGAAGACGATGATGAAGATGATGAAGAAGATGATGAAGATGAAGATGATGA  395

Qy  555  tacatttaaagcaaaagcacaagaaacttcgcgcgacagcttaatttatgatattacc  614
         |||| || |||  |||| | |||| | |   | ||| | ||| || | |||| ||
Db  396  TGAAGATGATGAAGAACGATGAAGACGACGAAATATGAATTAGAAGATGATGAAGATGA  455

Qy  615  aatgaaaagcgcgcgaagtacaagacgctgtgaaaagaagcagcgcaattttagacaaa  674
         | ||| | || || ||| |||||| |  || | ||||||| || | |  ||| ||||
Db  456  AGAAGATGATGAAGAACGATGAAGAAGAAGATGATGAAGATGATGAAGATGAAGATGA  515

Qy  675  tgtctgttgatcaaatcaatacttaccacaaaagtaacacagagatgctttcaaaactgaact  734
         | ||  ||  |   | ||  | || | ||||||| | ||  |  | || | ||||||| |||
Db  456  AGAAGATGATGAAGAACGATGAAGAAGAAGATGATGAAGATGATGAAGATGAAGATGAAGA  515
```

RESULT 2

```
LOCUS       AZ531291/c     877 bp    DNA      GSS        03-NOV-2000
DEFINITION  ENTBQ34TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ531291
VERSION     AZ531291.1   GI:11085838
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
ORGANISM    Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 877)
AUTHORS     Loftus,B., Van Aken,S. and Fraser,C.
TITLE       Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
JOURNAL     Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            The Institute for Genomic Research
            Department of Eukaryotic Genomics
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start: 22
            High quality sequence stop: 829.
FEATURES             Location/Qualifiers
     source          1..877
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
                     method for isolate identification. Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library construction is described in detail in Smith,
                     H.O. and Venter, J.C. (Making small insert libraries for
                     whole genome shotgun sequencing projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT    156 a    268 c     29 g    424 t
ORIGIN
```

Query Match   1.8%;   Score 68.2;   DB 245;   Length 877;
Best Local Similarity   45.0%;   Pred. No. 2.4e-05;

Matches 299;   Conservative 0;   Mismatches 363;   Indels 3;   Gaps 1;

```
Db  516 AGATGATGATGATGAAGAAGAAGATGATGATGAAGAAGAAGATGATGATGAAGAAGA 575
Qy  735 aacagaaagtaagcagaaagcattagatgatgaagtgcgcttactccaaagttga 794
Db  576 AGATGAAGAAGAAGAGATGATGAAGAAGAAGATGATGATGAAGAAGAAGATGATGAAGA 635
Qy  795 aagtgtaagtgcgattaacactcaaaacaaagctgttgaattaaccagcagtaccagtgaa 854
Db  636 CGACGAATATGAAATTAGAAGATGATGAAGAAGAAGATGATGAAGAAGAAGATGATGAAGA 695
Qy  855 cggaacactaaaattacaaacttcagctgctgaaatgaagataccaagtaaacgtaaatac 914
Db  696 TGATGAAGAAGAAGATGATGATGATGAAGAAGAAGATGATGATGAAGAAGAAGATGATGA 755
Qy  915 tgtacgtatctataaagtggacgtaa 941
Db  756 AGAAGATGAAGAGATGATGAAGATGATGA 782
```

RESULT 3

```
LOCUS       AZ551092       912 bp    DNA      GSS        14-NOV-2000
DEFINITION  ENTFJ22TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ551092
VERSION     AZ551092.1   GI:11176393
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
ORGANISM    Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 912)
AUTHORS     Loftus,B., Van Aken,S. and Fraser,C.
TITLE       Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
JOURNAL     Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            The Institute for Genomic Research
            Department of Eukaryotic Genomics
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
```

```
Qy  108 agtagtaagccaagcaaagcacagttcaaaaaagcacactacttacctaacaagt 167
Db  825 ATTAGAAGACCAAGATGAAGAAGAAGCCAAATATTGAAATTAGAAGAAGAAGATGATGAAGA 766
Qy  168 aacggaaactggtgaattcccaaacattaacgatgtatatgtgaatacacaacaaagcgaa 227
Db  765 TGAAGAAGAAGAAGATGATGATGAAGATGATGAAGAAGAAGATGATGAAGAAGAAGATGATGA 706
Qy  228 aaacgataccgttgatgcggtagcattagtgaataaagcaggtggcgcgaaaaaagcgc 287
Db  705 TGATGAAGAAGAAGATGATGATGATGAAGAAGAAGATGATGAAGAAGAAGATGA 646
Qy  288 ttacttagctgattacttcgctctacatttaaagcaacactcgcga 347
Db  645 TGATGAAGAAGAAGATGATGATGAAGAAGAAGAAGACGAATATGA 586
Qy  348 tggcgaagctcgttgtagcaacttacatcgatgcttacaactatgcaacaaattagacga 407
Db  585 TGATGAAGAAGAAGATGATGATGAAGAAGAAGATGAACGACGAATATGA 526
Qy  408 aatgcgccaagagctagaggctgctgttcaagcaaagatttagaaaaagcagaacaata 467
Db  525 ATTAG---AAGATGATGATGAAGAAGAAGATGAAGAAGATGATGAAGAAGAAGAAGATGA 469
Qy  468 ctatcacaaaattccttatgaaattaaaactcgacagtcaattttagatcgcgtatgg 527
Db  468 TGATGAAGATGATGATGAAGAAGAAGATGAAGAAGATGAAGAAGAAGAAGATGA 409
Qy  528 taaacaactcgtgattcgctctacatttaaagcacaaggaacttcgcga 587
Db  408 TGATGAAGAAGAAGATGATGATGAAGAAGAAGATGAACGACGAATATGA 349
Qy  588 cagcttaattatgtattaccgttgaaatgaaagcgcgaagtacaagacgctgtgaa 647
Db  348 ATTAGAAGATGATGATGAAGAAGAAGATGAAGAAGATGATGAAGAAGAAGATGA 289
Qy  648 agcaggcaatttagacaaagctgctgttgatcaaatcaatcaatacttaccaaa 707
Db  288 AGATGATGATGAAGAAGAAGATGAAGATGATGAAGAAGAAGATGATGATGA 229
Qy  708 agtaacaagatgcttccaaaactgaactaacacaagagtagcgaaaaagcattagtgcaga 767
Db  228 AGAAGAAGATGATGAAGAAGAAGAAGATGATGAAGAAGATGATGAAGATGATGA 169
Qy  768 tgaag 772
Db  168 AGAAG 164
```

Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 17
High quality sequence stop: 861.
Location/Qualifiers
1..912

FEATURES
source
/organism="Entamoeba histolytica"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="pHOS1; Site_1; Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + 1 method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."

BASE COUNT   423 a   34 c   264 g   191 t
ORIGIN

Query Match
Best Local Similarity   44.3%;
Matches 329; Conservative   0; Mismatches 408;   Indels   6; Gaps   1;

```
Qy   108 agtagtaaagccaagcaaaagcaacagttcaaaaaagcatactacttacagccatacagt 167
         ||||   | |||    ||||    |    |   |    || |  | |  |  || |  |||
Db   173 ATTAGAACGAGGAGCAAGAGACGAAGACGAATATGAATTAGAAGAAGAGATGATGATGA 232
Qy   168 aacggaaactggtgaattcccaaacattaacgatgtatatgctggatacaacaaaagcaa 227
         |||    | |      |      ||  ||  |  | | || |    ||       |  ||
Db   233 TGAAGAAGAAGAAGAGATGATGATGAAGATGAAGAAGAAGAGATGATGATGAAGAATGA 292
Qy   228 aaaacgataccgtgatgcggtagcattagtggccgaaaaagacgc 287
         |||     |  | ||    | |||  | |     |  |   ||
Db   293 TGATGAAGAAGATGATGATGAAGAAGATGATGAAGAAGAGANGAGA 352
Qy   288 ttacttagctgattacaaaagaatatgaaacttacgtttcaaagcaaacctaaatc 347
         ||| |  |   | | |  ||   ||    |     | |     ||  | |  | |
Db   353 TGATGAAGATGAAGATGATGAAGACGAATATGAATTAGAAGATGATGA 412
Qy   348 tggcggaagctccgtgtagcaacttacatcgatgcttacaacaaaattagacga 407
         || |  |   ||  | |   |  || | || | || |    |  |   |  |||
Db   413 TGATGAAGAAGAAGAAGATGATGATGAAGAAGACGACGAACGAGAATATGA 472
Qy   408 aatgcgccaagagctagaagggctgctgttcaagcaaaagtttagaaaaagcaagaacaata 467
         |||    |  |  | ||   | |  |  || | ||     | |    |    |  | |
Db   473 ATT-----AGAAGATGATGATGAAGAAGATGATGAAGATGATGAAGAAGATGAAGAAGA 526
Qy   468 ctatcacaaaattccttatgaaatctaaactcgcacagtcatttagatccgtatatgg 527
         || |   |    ||     ||     | |   |  | ||  |  | | |   | |
Db   527 TGATGATGAAGATGATGATGAAGAAGATGATGAAGAAGATGAAGATGAAGA 586
Qy   528 taaaacaactcgtgttacttcgctctacatttaaagcaaagcacagaacttcgca 587
         ||      | | | | | |  |   ||    | |  ||    | |  |  | |||
Db   587 TGATGATGAAGAAGAAGAAGATGATGAAGAAGAGACGATGAAGACGACGAATA 646
Qy   588 cagcttaattatgatattaccgttgaatgaaagacgcgcgaagtacaagacgctgtgaa 647
         ||||      ||    | | ||      |   |||   |    | |    | ||
Db   647 TGAATTAGAAGATGATGAAGATGAAGAAGATCATGATGAAGAAGAAGATGATGA 706
Qy   648 agcaggcaattagcacaaagctgctgttgatcaaatcaatcaataactattaccaa 707
         |||              |    | ||  |   |    |  |  |  | |  |||
Db   707 TGAAGATGATGAAGAAGACGAGATGAAGAAGATGATGAAGAAGATGAAGATGA 766
```

RESULT   4
AZ535744/c

LOCUS   AZ535744   816 bp   DNA   GSS   03-NOV-2000
DEFINITION ENTCQ25TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ535744
VERSION   AZ535744.1   GI:11092691
KEYWORDS   GSS.
SOURCE   Entamoeba histolytica.
ORGANISM   Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 816)
AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
TITLE   Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
JOURNAL   Unpublished (2000)
COMMENT   Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start:36
            High quality sequence stop: 816.
            Location/Qualifiers
            1..816

FEATURES
source
/organism="Entamoeba histolytica"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="pHOS1; Site_1; Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + 1 method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."

BASE COUNT   125 a   265 c   54 g   372 t
ORIGIN

Query Match
Best Local Similarity   43.9%;
Matches 287; Conservative   0; Mismatches 367;   Indels   0; Gaps   0;

```
Qy   708 agtaacagatgtcttcaaaactgaactaacagaagtagtagcaaaaagcattagatgcgag 767
         |||   | |  | || ||||     ||| | |||   |   | ||    ||||  || ||
Db   767 TGAAGAAGAAGATGATGAAGAAGAAGATGAAGAAGAAGAAGAAGATGAAGAAGATGA 826
Qy   768 tgaagctgcgcttactccaaaagttgaaagtgaagtgcgattaacactcaaaacaaagc 827
         ||||  | |  || ||||     | ||   |||  | | |||||   | |   |  |||
Db   827 TGAAGAAGAAGAAGAAGAGATGATGATGAAGACGACGAATATGAATTTAAGATGAGATGAAGAGA 886
Qy   828 tgttgaattaacagcagtaccag 850
         ||    ||  | |    |  |||
Db   887 TGATGAAGAAGATGATGAAGAAG 909
```

Query Match
Best Local Similarity   43.8%;
Matches 329; Conservative   0; Mismatches 408;   Indels   6; Gaps   1;

Score 68.2; DB 245; Length 912;
Pred. No. 2.4e-05;

Score 66.8; DB 245; Length 816;
Pred. No. 5e-05;

```
Qy   180 tgaattcccaaacattaacgatgtatatgctggatacaacaaagcgaaaaaacgataccg 239
         ||||| | | |   ||  | ||  | | || |    ||      |     |   | |||
Db   754 TGAATTAGAAGACGACGATGCATGATGAAGATGAAGATGATGAAGATGATGAAGATGA 695
```

```
LOCUS       AZ551618      843 bp    DNA       GSS    14-NOV-2000
DEFINITION  ENTDV54TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ551618
VERSION     AZ551618.1  GI:11176919
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica.
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 843)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: bjlofus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Class: shotgun
            Seq primer: M13-Reverse
            High quality sequence start: 39
            High quality sequence stop: 838.
FEATURES             Location/Qualifiers
     source          1..843
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site 1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S., (1993) Entamoeba histolytica: a
                     method for isolate identification. Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library construction is described in detail in Smith,
                     H.O. and Venter, J.C. (Making small insert libraries for
                     whole genome shotgun sequencing projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      155 a    266 c     18 g    404 t
ORIGIN
```

```
RESULT  5
AZ551618/c

Query Match  1.8%;   Score 66;   DB 245;   Length 843;
Best Local Similarity  43.7%;                            Pred. No. 7.7e-05;
Matches  340;  Conservative  0;  Mismatches  435;  Indels  3;  Gaps  1;

Qy   240 tgatgcggtagtagtgataaagcaggtgcgcggaaaagacgcttacttga 299
         ||||| |||| | ||||| | ||| |||||  |||||| |  |||||| |
Db   694 TGATGAAGAAGATGATGATGAAGAGATGATGATGAAGAAGAAGATGATGA 635

Qy   300 tttacaaaaagaatatgaaacttacgttttcaaagcaaacctaaatctggcgaagctcg 359
         ||| ||||| ||   |||  || |||| || |||  ||||  |   ||||  ||||| ||
Db   634 TGATGAAGAAGACGATGATGATGATGATGAAGAAGAAGATGATGATGAAGAAGAAGACGA 575

Qy   360 tgtagcaacttacatcgatgcttcaacactatgcaacaaaattagacgaaatgcgccaaga 419
         |  ||||| | |  | || |   | |||  ||||  ||||| || ||  ||||| || ||
Db   574 TGAAGAAGACGATGATGATGAAGAAGATGATGATGATGAAGAGATGATGAAGAAGAAGCGA 515

Qy   420 gctagaggctgtcttcaagcaaaagatttagaaaaagacagaacaatactatcacaaaat 479
         |||| || |  | || |||  ||||  || || |||||| || ||    |  |||| |||
Db   514 TGAAGAAGACGATGATGAAGAAGAAGATGATGATGAAGAAGACGATGAAGAAGAAGACGA 455

Qy   480 tccttatgaaattaaactcgcacagtcattttagatcgcgtatatgttaaaacactcg 539
         |||  | |||| ||| || | || ||  |  |||| |  | | || |||| |||  ||
Db   454 CGATGATGAATTAGATTAGAAGACGATGAAGATGAAGATGAAGAAGACGATGAAGAGTGA 395

Qy   540 tgattacttcgctcacattaaagcaaaagacacaagaacttcgcgacagcttaatta 599
         |||| | || |  | ||| |||  ||||| | ||| || |  || |||| || ||||
Db   394 TGAAGAAGACCATGATGATGAAGAAGAAGATGATGATGAAGAAGACGATGAAGAAGACGATGA 335

Qy   600 tgtattaccgttcgcaatgaaagacgcgcgaagtacaagacggctgttgaaagcagcaattt 659
         |  |||| | |  |  | ||||| | || ||| | ||||| ||  ||||||  |||||  ||
Db   334 TGAAGAAGACGATGATGATGAAGAAGACGATGATGATGAAGAAGACGATGATGAAGAAGACGATGA 275

Qy   660 agacaaagctaaagctgctgttgatcaaatcaatcaactaccaccaaagtaacagatgc 719
         || |||||   ||| |  | || || ||  |  ||  | || ||  |||  ||||||  |
Db   274 TGAAGAAGACGATGATGATGAAGAAGACGATGATGAAGAAGACGATGAAGAAGAACGATGA 215

Qy   720 tttcaaaactgaactaacagaagtagccagaaaaagcattagatgatgacgctcgcgct 779
         |||  ||||  |||  |||| |||   |||  |||| ||| |||| ||||| | | |||
Db   214 TGAAGAAGACGATGATGATGAAGAAGACGATGATGATGAAGAAGACGATGAAGAAGACGA 155

Qy   780 tacttcaaaagttgaaagtgcgattaacactcaaaacaaagctgttga 833
         ||| | |||| |||| || | || || ||   ||| ||||||  |||
Db   154 TGAAGAAGACGATGACGATGAAGAAGACGACGATGAAGAACGATGAAGA 101
```

```
Best Local Similarity  43.7%;

Qy   119 aagcaaaagcacagttcaaaaagcaatactattacttacagccattacagcggaactg 178
         || ||||| | ||| | |||  || |  | ||| | ||| || |  |  ||| ||||
Db   816 ATGAAGAAGATGATGATGAAGAAGACGATGATGATGATGAAGATGATGATGAAGAGATG 757

Qy   179 gtgaattcccaaacattaacgatgtatatgctgaataacaaagcgaaaaaacgatacc 238
         |  ||||  ||||| | ||| | |  | || | |||  |||| |  ||| | ||| |
Db   756 AAGAAGAGTGATGATGATGATGAAGAAGACGATGATGAAGAAGACGATGAAGAATAGAAT 697

Qy   239 gtgatgcggtagcattagtgaataaagcaggtggcgcggaaaaagacgcttacttagctg 298
         ||||| | || |  | ||| |   |||| || | |||| || ||| | ||  | ||||
Db   696 TAGAAGATGATGATGATGATGAAGAAGACGATGATGATGAAGAAGACGACCAAATAGTG 637

Qy   299 atttacaaaaagaatatgaaacttacgtttcaaagcaaacccaaatctggcgaagctc 358
         ||||||||| || | |||| ||  |||| | ||| ||||  ||||| |||| ||||
Db   636 ACGCAGATAATGAATTGAAGAAGATGATGATGAAGAAGAAGACGATGATGAAGAAGAAG 577

Qy   359 gtgtagcaacttacatcgatgcttcaacactatgcaacaaaattagacgaaatgcgccaag 418
         || |||  | | || | || |   |  ||  |||| ||||| |||| ||  ||| || ||
Db   576 AAGAAGATGATGATGATGATGAAGATGATGATGAAGAAGATGATGATGATGAAGAAGATG 517

Qy   419 agctagaggctgtctgttcaagcaaaagatttagaaaaagcaagaacaatgatgaagagatg 478
         ||||| ||  | |   | |||  ||||  || || || |  ||||| |||||||||| |||
Db   516 ATGAAGATGATGATGATGAAGAAGAAGATGATGATGAAGAAGAAGATGATGATGAAGAGATG 457

Qy   479 ttccttatgaaattaaactcgcacagtcgatatcgatatggtaaacaactc 538
         ||   | |||| |  ||| | | || |  |  | || |  | |||| ||
Db   456 ACGAATATGAATTTGAAGAAGATGATGATGAAGATGATGATGAAGAACACTG 397

Qy   539 gtgattacttcgctcacattaaagcaaaagaacaagaacttcgcgacagctaattt 598
         || || | ||  |  |||| |||  ||||| || ||| || || ||| ||  |||
Db   396 ATGAT---GATGAAGATGATGAAGAAGAAGATGATGATGATGAAGACGATGAAGAGATG 340

Qy   599 atgattattaccgttgcaatgaaagacgcgcgaagtacaagacgacgctgtgaaagcagcaatt 658
         | ||  ||| || |  | | |||| | || ||| | ||||| || | |||||||  |||||||
Db   339 AAGAGATGATGATGATGATGATGAAGAAGACGATGATGATGAAGAAGACGATGATGAAGAAGATG 280

Qy   659 tagacaaagctaaagctgctgttgatcaaatcaatcaatacttaccaaaagtaacagatg 718
         |||| ||||   |||  | | || || ||  |  ||  || |  ||  ||| ||||||
Db   279 AAGAAGAAGACGATGATGATGATGAAGATGATGAAGAAGATGATGATGAAGATGATGAAGATG 220

Qy   719 ctttccaaaactgaactaacagaagtagcgaaaaaagcattagatgatgatgctcgcgc 778
         | |||  ||||  |||  ||||  || || || || |  || ||||||||| | | |
Db   219 ATGAAGAAGATGATGATGATGAAGAAGACGATGATGATGAAGAAGACGATGAAGAAGAAG 160

Qy   779 ttactccaaaagttgaaagtgtcgattaacactcaaaacaaagctgttgaatta 838
         ||| | ||||| ||||  || | |||| || ||   ||| ||||  ||  ||
Db   159 ATGATGATGATGAAGAAGAAGACGATGATGAAGATGATGAAGAAGACGATGATGAAGAAGAAG 100
```

RESULT 6

Qy   839 cagcagtaccagtgaacggaacactaaaattacaacttcagctgctgcaaatgaaga 896
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   99 ATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGACGACGA 42

LOCUS   AZ529191   880 bp   DNA   GSS   03-NOV-2000
DEFINITION ENTBV68TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ529191
VERSION   AZ529191.1 GI:11081835
KEYWORDS   GSS.
SOURCE   Entamoeba histolytica.
  ORGANISM Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 880)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE   Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start: 17
            High quality sequence stop: 673.
FEATURES   Location/Qualifiers
     source   1..880
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a method described by Clark and Diamond (Clark,
                     C.G. and Diamond, L.S (1993) Entamoeba histolytica: a
                     method for isolate identification. Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + 1 method used for
                     the library construction is described in detail in Smith,
                     H.O. and Venter, J.C. (Making small insert libraries for
                     whole genome shotgun sequencing projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT   271 c   67 g   413 t   129 a
ORIGIN

Query Match   1.7%; Score 63.6; DB 245; Length 880;
Best Local Similarity 43.7%;
Matches 279; Conservative 0; Mismatches 359; Indels 0; Gaps 0;

Qy   49 gcaagtgcattgtcgctgcaaatccaaacgcttcgaaggcggctacagatgtagcaaca 108
          ||||| ||||||||| ||||| ||||||||| || |||| ||||||| |||||||
Db   750 GCAGAAGAAGAAGAGACGAAGAAGACGAAGAAGACGAAGAAGACGAAGATGAACAAGAAGA 691

Qy   109 gtagtaagcaagcaagcacagttcaaaaagcatactatacttacacgccataacagta 168
          |||||||| ||||| ||| |||||||| ||| ||| |||||| ||| |||||
Db   690 GCAGACGAAGATGACCAAGAAGAAGACGAAGAAGACGAAGAAGACGAAGAAGACGAA 631

Qy   169 acggaaaactgggtgaattcccaaacattaacgatgtatatgctgaatacaacaaagcgaaa 228
          ||||| |||||||| ||||| ||||||||||| |||||||||||||||| ||||||| ||
Db   630 GACGAAGAAGAAGACGAAGAAGATGAAGAAGACGAAGAAGACGAAGAAGAAGATGAA 571

RESULT 7

Qy   229 aaacgatacgtgatgcggtagcactagtgaataaaagcaggtggcgtggcgcaagaacgct 288
          |||||||| ||| | |||||| ||| ||| ||||| |||| | |||||| |||||||
Db   570 GACGAAGAAGATGAAGAGACGAAGAAGACGAAGAAGACGAAGAAGAAGACGAAGATGAA 511

Qy   289 tacttagctgatttacaaagagaatatgaaacttacgtttcaaagcaaaccctaaatct 348
          |||||||| |||||| ||| |||||||| ||||| ||| |||| |||||||||||
Db   510 GACGAAGATGATGAAGAGACGAAGAAGACGAAGAAGACGAAGAAGATGAAGAAGATGAA 451

Qy   349 ggcgaagtcgtgttagcacattacacatcgatgctctacaactcaacaaaattagacga 408
          |||| ||||| ||| ||| |||| | | |||||| || ||||||| ||||||||||
Db   450 GACGAAGATGATGAAGAGACGAAGAAGACGAAGATGAAGAAGACGAAGAAGATGAA 391

Qy   409 atgcgccaagagcctagaggctgctgttccaagcaaaagattagaaaaagcagaacatac 468
          ||||||| | | ||| ||| ||| |||||| ||| |||||| ||||| ||||| ||
Db   390 GACGAAGATGATGAAGACGAAGACGAAGAAGATGAAGAAGACGAAGATGATGAAGA 331

Qy   469 tatcacaaaattccttatgaaattaaaactcgcacagtcattttagtcgtgtatggt 528
          ||| |||||| |||| |||| |||||| || || ||| |||| ||||||||
Db   330 GACGAAGATGATGAAGACGAAGAAGACGAAGAAGACGAAGAAGATGAAGATGAA 271

Qy   529 aaaacaaactcgtgatttacttcgctctacatttaaagcaaaagcacaagaaacttcgcgac 588
          ||||| |||| |||| |||| || ||| ||| |||||||||||| ||| ||||| ||
Db   270 GACGAAGATGATGAAGAGACGAAGAAGACGAAGATGAAGAAGACGAAGATGAAGATGAA 211

Qy   589 agcttaatttatgatattaccgttgcaatgaaagcgcggaagtacaagaagctgtgaaa 648
          |||| |||| || |||| || |||||||||| | ||| || |||||||| ||||
Db   210 GACGAAGATGATGAAGAGACGAAGATGATGAAGACGAAGATGATGAAGACGAAGATGAT 151

Qy   649 gcaggcaatttagacaaagctaaagctgctgtgatca 686
          |||| ||||| || |||| ||||| ||| ||| |||
Db   150 GATGATGAAGACGAAGATGATGAAGAGACGAAGAAGATGATGA 113

RESULT 7

LOCUS   AZ546009   849 bp   DNA   GSS   14-NOV-2000
DEFINITION ENTFW53TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ546009
VERSION   AZ546009.1 GI:11167130
KEYWORDS   GSS.
SOURCE   Entamoeba histolytica.
  ORGANISM Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 849)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE   Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 26
            High quality sequence stop: 796.
FEATURES   Location/Qualifiers
     source   1..849
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a method described by Clark and Diamond (Clark,
                     C.G. and Diamond, L.S. (1993) Entamoeba histolytica: a

```
RESULT    8
LOCUS       AZ683582        891 bp    DNA         GSS         14-DEC-2000
DEFINITION  ENTKK47TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ683582
VERSION     AZ683582.1  GI:11820728
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 891)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 0208
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start: 16
            High quality sequence stop: 694.
FEATURES             Location/Qualifiers
     source          1..891
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a method described by Clark and Diamond (Clark,
                     C.G. and Diamond, L.S. (1993) Entamoeba histolytica: a
                     method for isolate identification. Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library construction is described in detail in Smith,
                     H.O. and Venter, J.C. (Making small insert libraries for
                     whole genome shotgun sequencing projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      400 a    55 c    291 g    145 t
ORIGIN
```

Query Match   1.7%; Score 61; DB 248; Length 891;
Best Local Similarity 43.5%; Pred. No. 0.0011;
Matches 277; Conservative 0; Mismatches 360; Indels 0; Gaps 0;

```
Qy    197 acgatgtatatgtgaatacaacaaagcgaaaaacgataccgtgatgcggtagcattag 256
          | |||||||||  | |||  | |  ||| ||  | ||||||  ||| | ||||| |||
Db     63 ATGATGAATTTGAATTAGAACGATGATGAAGACGATGATGAAG 122

Qy    257 tgaataaagccaggtggcgcgaaaaagaacgcttacttagctgattttacacaaaaagaatatg 316
          || | |  |  ||   |   |   || ||| | |||  | ||| |||   |  ||||||||
Db    123 ATGAGCGATGATGAAGACGATGATGAAGACGATGATGAAG 182

Qy    317 aaacttacgtttcaaacccctaaatctggcgaagctcgtgtagcaacttacatcg 376
          |  ||| || | |  |||    ||| |||  ||  | || | |||| | ||||
Db    183 AACAAGATGATGATGAAGACGATGATGAAGACGATGATGAAGACGATG 242

Qy    377 atgcttacaactatgcaacaaaattagaacgaaatgcgcaaagagctagaggctggttc 436
          |||  | |  |  | |   || |  ||| ||| || | |||||| | | ||| || |
Db    243 ATGATGACCATGAAGACGATGAAGACGATGAAGACGATGAAGACGATG 302

Qy    437 aagcaaaagatttagaaagaacaatactaccaaaattcattgaaattaaa 496
          |||   | |  |  ||| ||| | | | | | | ||| | ||| ||||||
Db    303 AAGAAGACCATGAAGACGATGAAGACCATGAAGACGATGAAGACG 362

Qy    497 ctcgcacagtcattttagatcgtataggtaaaacaactcgtgatttcgtctcta 556
          |  | ||| | | || ||||| |  |  |||| | || ||| | ||| | |  |
Db    363 ATGATGAATTTGAATTAGAACGATCATGAAGACGATGAAGACGATGAAGACGATG 422

Qy    557 catttaaagcaaagcaacaagaactccgcgacagcgcttaatttatgatattaccgtgtgcaa 616
          |  ||||||||||| ||||||| || | | |  |||  ||| |||| ||| ||| ||| ||||
Db         616
```

BASE COUNT   199 a   236 c   43 g   371 t

Query Match   1.7%; Score 62; DB 245; Length 849;
Best Local Similarity 44.2%; Pred. No. 0.00066;
Matches 303; Conservative 0; Mismatches 380; Indels 3; Gaps 1;

```
Qy     93 tacagatgtagcacagtagtaagcaaagcacagttcaaaaagcatactatac 152
          | | |   |  | || ||| | ||  | |||  |  ||||| ||| | ||||
Db    758 TGCCGAAGAAGATTCCTTTGACTTAGATGATGAAAAATGTAGAAACATAGATAATACCAGTT 699

Qy    153 ttacagccatacagtaacggaaactggtgaattcccaaacattaacgatgtatatgctga 212
          || | || | | ||| | || |  ||||| |  ||  |||  | | | ||| | ||| ||
Db    698 AGAAGAACAATGAATTAGAAGATGAATGATGAAATTGAATTAGAAGACGAAGATGA 639

Qy    213 atacaacaaagcgaaaaacgataccgtgatgcggtagcattagtgtgaataaagcaggtgg 272
          | |  |  |  ||   || |  |  |   |  |  |   || |||| | ||||| ||| ||
Db    638 TGATGATGATGAAGACCAAGAAGACGATGAAGACGATGATGAAGACGGAAGAGA 579

Qy    273 cgcgaaaaaagacgcttacttagctgatttaccaaaagaatatgaaacttacgtttcaa 332
          |  |||| |  | |  |  |||| ||  || || || ||  | |||||| | ||| | ||
Db    578 AGAAGAAGATGAATTTGAATTAGAAGACGAAGATGAAGACGATGATGATGATG 519

Qy    333 agcaaaacctaaatctggcgaagctcgtgtgcaacttacatcgatgcttacaactatgc 392
          || | |  ||||| |  |  |   ||| ||| |  ||||| | | |   || | |  ||
Db    518 AGAAGATGATGATGAAGATGAAGACGATGATGAAGATGATGATGAAGAAGA 459

Qy    393 aacaaaattagacgaaatgcgccaagagctagagggctgtctgttcaagcaaaagatttaga 452
          | |  |  |  | ||| |||  | |||| | |  | |||  |  ||||| ||||| |||| |
Db    458 CGAAGAAGAAGAGATGAATGATGAATTTGAATTAGAAGAGAATGATGAAGAAGATGA 399

Qy    453 aaaagcacgaacaactactatcacaaaattccttatgaaactcgcacagtccattt 512
          ||||||| | | |  | | | ||||| || ||| ||||| ||||||| || | |||
Db    458 AGAAGAAGAAGATGAATGAAGATGAATTTGAATTAGAAGAAGATGAAGAAGATGA 399

Qy    513 agatcgcgtatatgtaaaacaactcgtgattcgctccacatttaaagcaaagc 572
          | || || |  | |||||| | | | |   |||| | || | ||| ||||| ||
Db    338 AGATGCGAAGATGATGAAGATGAAGACGATGAAGACGATGATGAAGAGAGATGA 279

Qy    573 acaagaactt--cgcgacagcttaatttatgatattaccgttgcaatgaaagcgcgcga 629
          | ||| |||   |||| |  | |||| |||| ||| ||| |  | ||| ||| | | |||
Db    278 TGATGAATTTGAATTAGAACGATGAAGATGAATGATCATGATGATCAAGAGAAACGA 219

Qy    630 agtacaaagacgcgtgtgaaacgcaattgagacaaaagctaaacgctgtctgatcaaat 689
          || |   | |  || |  | |  | |||   | ||||  |  | | ||| | |||||||
Db    218 AGATGATGATGAATTTGAATTAGAAGATGAATGATGAATGATGAACGACGA 159

Qy    690 caatcaatacttaccaaaaagctaacagatgcttccaaaactgaactaacggaagtagcgaa 749
          |||| | |  ||| || ||| ||||| || | | |  ||| ||| | ||  ||| | | | |
Db    158 AGATGATGAAGAAGACGATGATGAAGAGAGATGATGAAGACGATGAAGACGATGA 99

Qy    750 aaaagcattagatgcgaatgaagctg 775
          || |  | |  ||| || ||||| ||
Db     98 AGATGATGATGAAGACGATGAAGAGATG 73
```

RESULT 9

```
Db   423 AAGAAGACGATGAGAAGACGATGAAGAAGACGATGATGAAGAAGACGATGAAGAAGACGATG
         ||    ||| ||||||   |||   ||||||||   |||||||| ||| |  ||||| ||
Qy   617 tgaaagcgccgaggtacaacgcgtgtgaaagcgcaattttagacaaagctaaagctg 676
Db   483 AAGAAGACGATGATGAAGAAGACGATGAAGAAGACGATGAAGAAGACGATGATGAAGAAG 542
         ||  |||||||||||||| ||||||||| ||| |  ||||||||| ||||||||||||
Qy   677 ctgttgatcaaatcaatcaatacttaccaaaaagtaacagatgcttcaaaactgaactaa 736
Db   543 AAGAAGACGATGATGAAGAAGACGATGATGAAGAAGACGATGATGAAGAAGACGATGATG 602
         ||||||||||||||| |||||| |||| |||||| ||||||||| ||||||||||| |
Qy   737 cagaaagtagcgaaaaagcattagaggatgagaagcgcgcttactccaaaagttgaaa 796
Db   603 AAGAAGATGAAGAATGAAGAAGATGAAGATGAAGAAGATGAAGAAGAAGATGAAGATGAAG 662
         |||||||||||||| |||||||||||||||| |||||||||||||||||  ||||| |
Qy   797 gtgtaagtgcgattaacactcaaaacaaagctgttga 833
Db   663 ACGACGATGAAGACGACGATGAAGAACGACGATGAAGA 699
```

LOCUS   AZ549980   900 bp   DNA   GSS   14-NOV-2000
DEFINITION   ENTDD94TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
             genomic, DNA sequence.
ACCESSION   AZ549980
VERSION   AZ549980.1 GI:11175122
KEYWORDS   GSS.
SOURCE   Entamoeba histolytica.
ORGANISM   Entamoeba histolytica
             Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 900)
AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
TITLE   Determination of clone end sequences from Entamoeba histolytica
             HM1:IMSS sheared DNA library
JOURNAL   Unpublished (2000)
COMMENT   Contact: Brendan J Loftus
             Department of Eukaryotic Genomics
             The Institute for Genomic Research
             9712 Medical Center Dr., Rockville, MD 20850, USA
             Tel: 301 838 0200
             Fax: 301 838 0208
             Email: bjlofus@tigr.org
             Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
             DNA library
             Seq primer: M13-Forward
             Class: shotgun
             High quality sequence start: 20
             High quality sequence stop: 890.
FEATURES   Location/Qualifiers
     source   1..900
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a method described by Clark and Diamond (Clark,
                     C.G. and Diamond, L.S. (1993) Entamoeba histolytica: a
                     method for isolate identification. Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + 1 method used for
                     the library construction is described in detail in Smith,
                     H.O. and Venter, J.C. (Making small insert libraries for
                     whole genome shotgun sequencing projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT   376 a   63 c   235 g   226 t
ORIGIN

Query Match

Best Local Similarity   44.8%;
Matches 231; Conservative   0; Mismatches 285; Indels   0; Gaps   0;
Pred. No. 0.002;

```
Qy   260 ataaagcaggtgcgccgaaaaagacgcttacttagctgatttacaaaagaatatgaa 319
         |||| |  ||    |||||||  |||| ||| |||||||  |||| |||||  |||||
Db   254 ATAATGACTATGAAGTTAGAAAAACAATGAAGAATTTAGATGACGAAGAATTTGACTTAGAAG 313
Qy   320 cttacgttttcaaagcaaacccctaaatctggcgaagcctcgtgtagcaactcacatcgatg 379
         || |||  |||||| |||| ||   |||| ||||| |||||| ||| | ||| ||| ||||
Db   314 ATGAAGATGAAGATGAAGAAGACGATGATGAAGACGATGAAGATGAAGAAGATGAAGATGAAG 373
Qy   380 cttacaactatgtcaaaaaattagacgaaatggccaaagagctagagggctgttcaag 439
         || ||||| |||  ||||||| || || ||  |||| || |||| |  || | |||||
Db   374 ACGATGAAGATGAAGAAGATGATGAAGATGAAGATGAAGAAGATGAAGAAGATGAAGATG 433
Qy   440 caaaagatttagaaaaagcagaacaataactatcccaaaaattccttatgaaattaaactc 499
         |||||||| ||| |||||| ||| |||| |   |||||||||| ||  ||||| ||||||
Db   434 AAGAAGACGATGAAGAAGACGATGAAGATGAAGAAGAAGACGATGATGAAGATGATGAAGATG 493
Qy   500 gcacagtcatttttagatcgcgtatatggtaaaacaactcgtgatttacttcgtctacat 559
         |  || || ||  ||||||| | |||   |  |||  ||||| | || ||||||| ||
Db   494 ATGAAGATGATGAAGAAGATGATGAAGATGAAGATGAAGATGAAGATGAAGAAGATCATGAAG 553
Qy   560 ttaaagcaaagcacacaagacacttcgcgacagcttaattatgatattaccgttgcaatga 619
         || ||||||| ||| |  |||| |||  |||| |||| |  |||| |||| |||| ||||
Db   554 ATGATGAAGACGATGAAGATGATGAAGATGAAGATGATGAAGATGAAGACGAAGATGAAGATG 613
Qy   620 aagccgcgggaagtacaagacgtcgtgtgaaagcaggcaaatttagacaaagctaaagctg 679
         |||   | ||| ||| ||||| | |||||||| || | |||||||||||||||||| ||
Db   614 AAGAAGATGAAGATGAAGAAGACGATGATGAAGACGATGAAGATGATGAAGATGAAGATG 673
Qy   680 ttgatcaaatcaatcaatacttaccaaaaagtaacagatgctttcaaaactgaactaacag 739
         |||| |  |||  | ||||| ||||| ||||  || || ||  ||| ||| |||| || |
Db   674 ATGAAGACGATGATGAAGATGATGAAGACGATGAAGACGATGATGAAGAAGACGATGAAGACGATG 733
Qy   740 aagtagcgaaaaagcattagatgcagatgaagctg 775
         |||| |  |||  || |||| | | || || ||||
Db   734 ATGAAGACGATGATGAAGATGAAGAAGACGATGAAG 769
```

RESULT 10

LOCUS   CNS0134P/c   997 bp   DNA   GSS   26-JUL-1999
DEFINITION   Drosophila melanogaster genome survey sequence SP6 end of BAC
             BACN09C07 of DrosBAC library from Drosophila melanogaster (fruit
             fly), genomic survey sequence.
ACCESSION   AL102403
VERSION   AL102403.1 GI:5614014
KEYWORDS   GSS.
SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
             Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
             Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
             Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 997)
AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
             BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
             - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
             collaboration with the European Drosophila Genome Project (EDGP) -
             http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
             library (Dros BAC) was made by Alain Billaud at CEPH (Centre
             d'Etude du Polymorphisme Humain) with funding provided by a MRC
             project grant. The DNA was prepared from embryos by Alain Bucheton
             and Genevieve Payan. It has been constructed in the vector
             pBeloBAC11.
FEATURES   Location/Qualifiers
     source   1..997
                     /organism="Drosophila melanogaster"
                     /plasmid="pBeloBAC11"
                     /db_xref="taxon:7227"

```
                                    /clone_lib="DrosBAC"
                                    /clone="BACN09C07"
                                    /note="end : SP6"
                       V_type: pBluescript II KS
                       PRIMER: KS
```

Query Match 1.68; Score 59.8; DB 219; Length 997;
Best Local Similarity 37.5%; Pred. No. 0.0022;
Matches 245; Conservative 43; Mismatches 362; Indels 3; Gaps 1;

BASE COUNT   162 a   115 c   79 g   124 t

ORIGIN

```
Qy   437 aagcaaaagattagaaaagcagaacaatactatccaaaaattccttatgaaattaaa  496
         |||:|| |||| ...
Db   656 AAAAAAAWII...                                               597

Qy   497 ctcgcacagtcattttagatccgtatatggtaaaacaactcgtgatttacttcgctcta  556
Db   596 ATTAAAAAATAAAA...WWGAAAGRAGAARAAAAGGAAARARARAAWAWAAAAA        537

Qy   557 catttaaagcaaaagcacaagaacctcgcgacagcttaatttatgatattaccgttgcaa  616
Db   536 WAAWAAAATTRAAAWRTRAAAWAAAGAAATAWAAARTRAAAAGAAWAWAAAAAGGWA     477

Qy   617 tgaaagcgcgcgaagtacaagacgctgtgaaagcagaggcaatttagacaaaagctaaagctg  676
Db   476 AATAAAAAATAAAWAAAAAGATAAAAAAWGAAAAAAAAWTRAAAAAAATAAAAWAAAA    417

Qy   677 ctgttgatcaaatcaatcaatacttaccaaaaagtaacagatgcttcaaaactgaactaa  736
Db   416 WTGAWAAAAGWA...AAAATTAGWAGRATPARATAARWTGAAATAA               357

Qy   737 cagaagtagcgaaaaaaagcattagatgcagatgaagctgcgcttactccaaaagttgaaa  796
Db   356 AAGAAAARAAARA...TWAAAWRAAAACATAATRWGNGTAAAACAAAAAAAGAAAA      297

Qy   797 gtgtaagtgcgattaacactcaaaaacaaagctgttgaattaacagcagtaccagtgaacg  856
Db   296 AAAAARAARARAWARAWAWNAAAAGARAAAANNNNNNNANANNNNNNANNNNN         237

Qy   857 gaacactaaaattacacttcagcgtgctgcaaatgaagatacagtaaagtaaataactg  916
Db   236 NNGGANNNNANNNAAANANATGTTRCAATWAWAAGTATRGAAAAAATAARWAAAAAA--   178

Qy   917 tacgtatctcataaagtggacggtaacatcttccattgccttaataacggcagagtgttctt  976
Db   179 ARARARAAGWARAAAAWGAAWRTWTCCAAAGAAGAGGAGGTTTAWAAGTTAWAGACATGAMC  120

Qy   977 tatctacagacggaaaactatcactgtggatgcttcaactccattcgaaaataataccg  1036
Db   119 TTRAGCAAAATWAAGGWAATRAAAAWTRAAAAAWTAAAAAAGTTAACAAAAAACGGTAAWWTAA  60

Qy  1037 agtataaagtagttaaaggtattaaagacaaaaatggcaaaagaaattaaa  1089
Db    59 ACCGGAAACAAGTTGTAAAAANATGAAAAAAAATTATTNGMATTGWAATRAAM  7
```

RESULT 11
FR0048073
LOCUS   FR0048073   494 bp   DNA   GSS   05-JAN-2001
DEFINITION   Fugu rubripes GSS sequence, clone 263K15bD8, genomic survey
             sequence.
ACCESSION   FR0048073
VERSION   FR0048073.1 GI:12052694
KEYWORDS   GSS; genome survey sequence.
SOURCE   Takifugu rubripes
  ORGANISM   Takifugu rubripes
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
             Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
             Tetraodontidae; Takifugu.
REFERENCE   1 (bases 1 to 494)
  AUTHORS   Clark,M.S.
  TITLE   Direct Submission

  JOURNAL   Submitted (29-SEP-2000) MRC Human Genome Mapping Project Resource
             Centre Hinxton, Cambridge, CB10 1SB, UK Email:
             biohelp@hgmp.mrc.ac.uk
COMMENT   Vector: pBluescript II KS
             V_type: phagemid
             PRIMER: KS
             DESCR:
             One pass dye-terminator sequencing of BAC (pBeloBACII) cloned
             genomic sequence
             The BACs can be obtained from http://www.incyte.com.
FEATURES   Location/Qualifiers
     source   1..494
                     /organism="Takifugu rubripes"
                     /db_xref="taxon:31033"
                     /clone_lib="263K15bD8"
                     /clone="263K15bD8"
BASE COUNT   128 a   163 c   45 g   158 t

ORIGIN

Query Match 1.68; Score 58.8; DB 258; Length 494;
Best Local Similarity 48.0%; Pred. No. 0.0032;
Matches 168; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

```
Qy  1259 cgaacgttactatcagatgttgaaactggaaaacgcattccagtaattgcatctactt  1318
         ||||| ||||  |||  ||||  |||||   |||  |||||  ||||||
Db   145 CGACTGCTACTACTACTGCTACTACTACTACTACTACTNCTGCTACTACTA          204

Qy  1319 ctggttctacaaattactattacgttaaaagaagcgttagtaactggtaaacaaatataaac  1378
         ||||| ||| |||||  ||| ||||  ||||||  |||  ||||
Db   205 CTACTACTACTGCTACTACTACTACTACTACTACTGCTACTACTACTACTA          264

Qy  1379 ttgctcatcaatatgttaaaacattaaactggttacaatgcagaagcttacgagttagtgt  1438
         ||||| |||| |||||| |||||  ||||  ||||| |||
Db   265 CTACTACTACTACTACTACTGCTACTACTGCTACTACTACTACTACTACTA          324

Qy  1439 tcactgcaaacgcatcagcaccaactgtgttgctaccgccctcactcttaggtgttgtacaa  1498
         ||  ||||  ||||  ||||  |||  |||||| |||||| ||||||  ||||||
Db   325 CTACTACTACTACTACTGCTGCTGCTGCTACTACTACTACTACTACTGCTACTACTA       384

Qy  1499 ctttatctactgttctcttacaacaaaatgtttgggtaaatgtaaatggcttggtggttgtgaatg  1558
         |||||||||| ||||  |||||| |||  ||||| ||||  |||  ||||| |||||| ||||
Db   385 CTACTACTGCTACTACTACTACTACTACTACTGCTACTACTGCTACTACTACTACTG        444

Qy  1559 aagtggaacttattatcctggtctccaattcacaacaacgtttgctact  1608
         ||||  |||| |||| ||||||| |||| ||||||  |||| |||||||
Db   445 CTACTACTACTACTACTACTACTACTACTGCTACTACTACTACTGCTACTACT      494
```

RESULT 12
AZ550256/c
LOCUS   AZ550256   905 bp   DNA   GSS   14-NOV-2000
DEFINITION   ENTEV58TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
             genomic, DNA sequence.
ACCESSION   AZ550256
VERSION   AZ550256.1 GI:11175557
KEYWORDS   GSS.
SOURCE   Entamoeba histolytica.
  ORGANISM   Entamoeba histolytica
             Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 905)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE   Determination of clone end sequences from Entamoeba histolytica
             HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Brendan J Loftus
             Department of Eukaryotic Genomics
             The Institute for Genomic Research
             9712 Medical Center Dr., Rockville, MD 20850, USA
             Tel: 301 838 0200
             Fax: 301 838 0208
             Email: bjloftus@tigr.org
             Clones are derived from the Entamoeba histolytica HM1:IMSS sheared

RESULT 13
LOCUS   AZ527724   912 bp   DNA   GSS   03-NOV-2000
DEFINITION ENTBD71TR Entamoeba histolytica Sheared DNA Entamoeba histolytica

ACCESSION   AZ527724
VERSION   AZ527724.1 GI:11079851
KEYWORDS   GSS.
SOURCE   Entamoeba histolytica.
  ORGANISM Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 912)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE   Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Class: shotgun
            Seq primer: M13-Reverse
            High quality sequence start: 14
            High quality sequence stop: 825.
FEATURES   Location/Qualifiers
     source   1..912
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
                     method for isolate identification. Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + 1 method used for
                     the library construction is described in detail in Smith,
                     H.O. and Venter, J.C. (Making small insert libraries for
                     whole genome shotgun sequencing projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT   203 a   260 c   46 g   396 t
ORIGIN

Query Match   1.6%; Score 58.4; DB 245; Length 905;
Best Local Similarity   44.5%; Pred. No. 0.0046;
Matches 233; Conservative   0; Mismatches 291; Indels   0; Gaps   0;

```
Qy   172 gaactgttgaattcccaaacattaacgatgtatatgctgaatacacaaagcgaaaaa 231
         ||  |||| |     ||       ||    |  | |||||||   | ||| ||   ||
Db   726 GATGATGATGAAGACCGAAGCAAGAGATGATGAAGAAGAGATGATGACGAAGAAGAA 667

Qy   232 cgataccgtgatgcggtagcattagtgaataagcaggtgtgccgaaaaagacgcttac 291
         | |  || | |||  |   |  | |||  || ||  |  |  || |||| | | | | ||
Db   666 GAAGATGATGATGAATTGAATTAGAGATGATGAAGATGAACGAGATGATGATGAAGAA 607

Qy   292 ttagctgattacaaaagaatatgaaacttacgttttcaaagcaaacctaaatctggc 351
         |   | | |     ||   |  | | |||  |||||| |||| ||    |||| | ||
Db   606 GATGATGATGAAGAACGAGAAGATGATGAAGAAGAGATGATGAAGAACCGAA 547

Qy   352 gaagctcgtgtagcaacttacatcgatgctcacaacaaaattgacgaaatg 411
         ||||   |  | |     || | | ||   |   ||| | ||   |  |  |
Db   546 GAAGAAGATGATGATGATGATGAATTAGAGATGAAGAAGAGATGAAGAGA 487

Qy   412 cgccaagagctagagcgctgttcaagcaacaaagacagaacaatactat 471
         | |     || | | |    |||  || | ||   |  || || || | |
Db   486 GATGATGATGATGAAGACGAAGATGAAGAAGATGAAGAAGATGAAGAAGAT 427

Qy   472 cacaaaattccttatgaaattaaacctcgcacagtcatttttagatcgtgtatatgtaaa 531
         |  || |  |  | || ||| |   |  |   | | |  |||||| | | |  | | |  |
Db   486 GATGATGATGATGAAGACGAAGATGAAGAAGATGAAGAAGATGAAGAAGAT 427

Qy   532 acaactcgtgattacttcgtctctacattttaaagcaagacaacttcgcgacagc 591
         | |  |  || ||| |  ||| | | |   || ||||  ||   ||| | | | ||
Db   426 GATGAAGATGATGATGATGATGATGATGATGAAGATGATGATGATGAAGAA 367

Qy   592 ttaattattatgaatataccgttgcaatgaaagcgcgcgaagtacaaagctgtgaagca 651
         ||  ||| |||     | | | || | | | | ||  |  || | | | | |||   | |
Db   366 TTAGAGATGATGAAGATGAAGAGATGATGATGATGATGATGATGATGAA 307

Qy   652 ggcaattagacaaagctaaagctgctgttgatcaaataacatca 695
         ||  ||  | |   ||| |  ||  ||||| || || |    | |
Db   306 TTTGAATTAGAGATGATGAAGAAGATGATGAAGATGATGATGA 247

Qy   246 GAAGAAGATGATGATGATGATGAATGAAGAAGATGATGAAGATGATGA 203
         ||||||||||||||||||||||| |||||||||||||||||||||||||
Db   246 GAAGAAGATGATGATGATGATGAAGATGAAGAAGATGATGAAGATGATGA 203
```

---

genomic, DNA sequence.
ACCESSION   AZ527724
VERSION   AZ527724.1 GI:11079851
KEYWORDS   GSS.
SOURCE   Entamoeba histolytica.
  ORGANISM Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 912)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE   Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT   Contact: pHOS1
            Department of Eukaryotic Genomics
            Institute for Genomic Research (TIGR), Rockville, MD.
            Genomic DNA isolated from broth cultures of E. histolytica
            using a method described by Clark and Diamond (Clark,
            C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
            method for isolate identification. Exp. Parasitol.
            77:450.). The DNA was mechanically sheared to give a
            tight size distribution (~2 kb). The v + 1 method used for
            the library construction is described in detail in Smith,
            H.O. and Venter, J.C. (Making small insert libraries for
            whole genome shotgun sequencing projects. In Genome
            Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barell, Oxford University Press, 1999)."
FEATURES   Location/Qualifiers
     source   1..912
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Class: shotgun
                     Seq primer: M13-Reverse
                     High quality sequence start: 14
                     High quality sequence stop: 825.
BASE COUNT   455 a   97 c   149 g   211 t
ORIGIN

Query Match   1.6%; Score 57.4; DB 245; Length 912;
Best Local Similarity   45.3%; Pred. No. 0.008;
Matches 291; Conservative   0; Mismatches 346; Indels   6; Gaps   2;

```
Qy   170 cggaactggtgaattcccaaacattaacgatgtatatgctgaatacacaaagcgaaaa 229
         |  | ||| |    ||      ||    | ||    | |  ||    | || || ||   |
Db    72 CAGATAATGGAGAAGATGATGAATATTTGAAGAAGAGTAGAGATTTACATCGAGCAGCT 131

Qy   230 aacgatacgtgatgcggtagcattagtgaataagcaggtgtgccgaaaaaagacgctt 289
         ||  |  | || ||  |  | | | ||  || || | | || | | |  |||| | | |
Db   132 ATCAATCTAGTACTTATTTTGATCCAGACAAATGAACCAGACAGAGTAAATAAACCAGACA 191

Qy   290 acttagctgattacaaaaagaatatgaaacttacgttttcaaagcaaaccctaaatctg 349
         ||  || | |  |  ||||| |  |||   |||  |||||| ||||   ||| ||| ||
Db   192 TAATAGATGAAGAAATTAACAAAACAGAAGAGATTCAGCATTGATG---ATCCAGATATG 248

Qy   350 gcgaagctcgtgtagcaacttacatcgatgctcaacaactatgcaacaaaattagacgaaa 409
         |   |   | | | |    || | | ||   |   |||  |  ||| | | || |  | | |
Db   249 AAGAATATACTAAAGAAATACATATCGATGAAGAACACATTGAAGAAAAACAAAAGACAAG 308

Qy   410 tgcgccaagagagctgctgttcaagcaacaaagaacattagaaaagcagaacaatact 469
         |   |     |  || | ||  || |  ||   | || | ||||||| |  ||  | |
Db   309 AAACGGAAGAGAAGAGAAAAATGAATTACCTCTACAAACAATAGTTGATCCAGAAAATCCAA 368

Qy   470 atcacaaaattccttatgaaattaaaactcgcacagtcatttttagatcgtgtatatggta 529
         ||  ||  | |  | ||  ||| |   || ||||| || | |||| |  | |   |  |||
Db   309 AAACGGAAGAGAAGAGAAAAATGAATTACCTCTACAAACAATAGTTGATCCAGAAAATCCAA 368
```

RESULT 14

```
A2676392/c
LOCUS       A2676392      938 bp   DNA             GSS       14-DEC-2000
DEFINITION  ENTHY47TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   A2676392
VERSION     A2676392.1  GI:11813538
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
ORGANISM    Entamoeba histolytica.
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 938)
AUTHORS     Loftus,B., Van Aken,S. and Fraser,C.
TITLE       Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
JOURNAL     Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start: 34
            High quality sequence stop: 809.
FEATURES             Location/Qualifiers
     source          1..938
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a method described by Clark and Diamond (Clark,
                     C.G. and Diamond, L.S. (1993) Entamoeba histolytica: a
                     method for isolate identification. Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library construction is described in detail in Smith,
                     H.O. and Venter, J.C. (Making small insert libraries for
                     whole genome shotgun sequencing projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin and B.
                     Barell, Oxford University Press, 1999)."
```

```
Qy   369 ATAATATATTTATACATATGTAATAATAACTATATAGAATTATTAAAAGATATGAAAATAAATA
         |||  || ||||||||||  ||||    ||||||||
Db   369 ...
Qy   429 TTAACAATAAAAAGGAAATAATAAAAACGTTAACAAGATATACAGATTAGACAAA 428
Db   429 ...
Qy   489 ACAATTATGAGAATTTAAAAATGCAATAAAGAGATGATAGAAGTAATAATCCCTG 548
Db   489 ...
Qy   530 aaacaactcgtgattacttcgctcacatttaaagcaaagcacaagaacttcgcgaca 589
Qy   549 TAAATAGAAGAAGAGAGAGTGATATAATAGTTACAGAAATAGATTATGAAATAAAAG 608
Qy   590 gcttaatttatgatattaccgttcaatgaaagtacaagacgctgaaag 649
Qy   650 caggcaaatttagacaaagctaaagctgcttgatcaaatcaatacttaccaaag 709
Qy   609 ATAAACAAATAGGAAGAAAATACAATACAACAAGATG--AAAAAGAAACCAGATAGAAATA 665
Qy   666 TTAAACACCTAGACCACCAAGACAGAGAATAGAAAATTACTATTTA 708
Qy   709 caggcaaatttagacaaagctaaagctgcttgatcaaatcaatacttaccaaag 709
Qy   710 taacagatgcttcaaaactgaactaacagagagtagcggaaaaaagcattagatgatgcagatg 769
Qy   770 aagctgcgcttacctccaaaagtgaaagtgtaagtgcgattaa 812
```

```
Query Match     1.68;   Score 57.4;   DB 247;   Length 938;
Best Local Similarity   49.1%;   Pred. No. 0.008;
Matches  182;   Conservative   0;   Mismatches 186;   Indels   3;   Gaps   1;
```

```
Qy   561 taaagcaaaagcacaagaacttcgcgcacagcttaattatgatattaccgttgcaatgaa 520
Db   764 TGAAGCAAAAGCTAAAAGCTAAAGAAGAGCTGAAGCAAAAGCTAAAGAAGCTGAAGCAAAAGC 705
Qy   521 agcgcgcgaagtacaagacgctgaaagcaggcaaatttagacaaagctaaagctgctgt 680
Db   704 TAAAGAAGAAGCTGAAGCAAAAGCCAAAGAAGCTGAAGCAAAAGCTGAAGCAAAAGC 645
Qy   681 tgatcaaatcaatcaatacttcccaaaaagtaacagatgcttccaaaactgaactaacaga 740
Db   644 TGAAGCAAAAGCCAAAGAAGCTGAAGCAAAAGCTAAAGAAGCTGAAGCAAAAGC 585
Qy   741 agtagcgaaaaaagcattagatgatgcagatgaagctgcgcttacctccaaaagtgaacgaac 800
Db   584 TAAAGAAGAAGCTGAAGCAAAAGCTAAAGAAGAAGC---TGAAGCAAAAGCTAAAGAAGA 528
Qy   801 aagtgcgattaacatccaaaacaaaagctgttgaattaacagcagtaccagtgaacgaac 860
Db   527 AGCTGAAAAAGAAAAAGCTAAAAGCTAAAGAAGAAAATAAAGAAGAAGTTAAAGAAGA 468
Qy   861 actaaaaattacaacttcagctgctgcaaatgaagatacagtaaacgtaaatactgtacg 920
Db   467 AGTTAAAGAAGAAGTTAAAGAAGAAGTTAAAGAAGAAAATAAAGAAGAAGAGA 408
Qy   921 tattcataaag 931
Db   407 AAAATAAAGAAG 397
```

RESULT 15

```
A2550718/c
LOCUS       A2550718      879 bp   DNA             GSS       14-NOV-2000
DEFINITION  ENTEM36TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   A2550718
VERSION     A2550718.1  GI:1117601
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
ORGANISM    Entamoeba histolytica.
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 879)
AUTHORS     Loftus,B., Van Aken,S. and Fraser,C.
TITLE       Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
JOURNAL     Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 25
            High quality sequence stop: 801.
FEATURES             Location/Qualifiers
     source          1..879
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
```

Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + 1 method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."

BASE COUNT   193 a   219 c   72 g   395 t
ORIGIN

Query Match   1.5%;   Score 56.8; DB 245; Length 879;
Best Local Similarity   43.2%;   Pred. No. 0.011;
Matches 271; Conservative 0; Mismatches 357; Indels 0; Gaps 0;

```
Qy    85 gaagcggctacagagtgtagcacagtagtaagccaagacaaaagcacagttcaaaaagca 144
          | ||·| |·||·|  |·||·|||·||||·|·||·||||·|  || ·|| ||| || |||
Db   679 GTAGAGCTTTTTGAGAGAAGAAAATAATGATGATGAAGAATGAAATTCCAATTACGAGAGA 620

Qy   145 tactatacttacagcatacagtaacggaaactggtgaattcccaaacattaacgatgta 204
          | |·| ||| ||·|| ·|·|||·||·||·|||·|||·||·|||·|||·||·|||·||·||
Db   619 GAAGATGATGAAGAAGATGAAGAAGATGAAGAAGATGAAGAAGATGAAGAAGATGAAGAA 560

Qy   205 tatgctgaatacacaacaaaagcgaaaaaacgatacacgtgatgcggtagcattagtgaataaa 264
          |||·|    ||·| ·||·|·|·| ·|||·||·|·|||·||·|||·|·|||·||·||·||·|| |||
Db   559 GATGAAGAAGATGAAGAAGATGAAGAAGATGAAGAAGATGAAGAAGATGAAGAAGATGAA 500

Qy   265 gcaggtggcgaaaaaagacgcttacttagctgattacaaaaagaatatgaaacttac 324
          | | || |·|·|·|·|·|·||·||||||·|·||·|||·|||·|||·||·|||·|| |||
Db   499 GAAGATGAACAAGACAGAAGAAGAGATGAAGAAGATGAAGAAGATGAAGAAGAGAT 440

Qy   325 gtttccaaagcaaaaccctaaatctgtcgaagctcgtgtagcaacttacaatcgatgcttac 384
          |·| | ·|  |·| | | ·|·| ·|·||·|·|·||||·||·||·|·|||·||·|·|||·|||
Db   439 GAAGAAGATGAAGAAGATGAAGAAGACGAAGAAGAAGACGAAGAAGAAGAAGAGAA 380

Qy   385 aactatgcaacaaaattagacgaaatgccaaagaagctagaggcgtgctgttcaagcaaa 444
          |·|| | ·|| |·|||·||·||·| |·||·|·|·||·||||·||·||·|||·||·|||·|||
Db   379 GATGAAGAAGATGAAGAAGACGAAGAAGATGAAGAAGACGAAGAAGAACGAA 320

Qy   445 gatttagaaaaagcagaacaatactatcacaaaattccttatgaaattaaaactcgcaca 504
          |·|·||·| |·|·|·|·| ·|·| ·||| ·|·|||·||·|||·|·||·|||·|||·||·|| |
Db   319 GAAGACGAAGAAGACGAAGAAGAAGCGAATAATGATTAGAGGAAGAGAAGAGGATGATGAT 260

Qy   505 gtcattttagatcgcgtatatggtaaaacaactcgtgattactccgctctacattaa 564
          |·| ||·||·||·||·| |·|·||·||·|||·||·|·|||·||·|·||·|·|||·|·|||
Db   259 GCTGTTGCCGAGTCTTGTTGTTGCTGAAYTTGATGATGAGACCTTCTTGACTCTGAGTGAA 200

Qy   565 gcaaaagcaccagagaacttccgacagcttaatttatgatattaccgttgcaatgaaagcg 624
          ·| |·|·|·|·| ·||·|·||·||·|·|||·||·|·|||·|·|·| ·|·||·|| ||·|||·|
Db   199 GGTCTTGATGATGACGGATGAATTAGATCTTGATGATGACGAAGATTTAGAAGAAGAACCA 140

Qy   625 cgcgaagtacaaagacgctgtgaaagcaggcaattagacaaagctaaagctgctgttgat 684
          |·|·|·|·|·|·|·||·|·|·|||·|·|·| ||·||·||·|·|·|·|·|·|·|·|·|·|·|||
Db   139 GACTTTTTTGAAGACGAAGAGATGAACAAGTGATACTCAAGAAGAAGAGAAGAATATGAA 80

Qy   685 caaatcaatcaatacttaccaaaagtaa 712
          ·|·||·|·|·|·|·|·|||·|·|||·|||
Db    79 TTAACTCAAGAATGGTCAGAACAAGCAA 52
```

Search completed: July 4, 2001, 23:34:07
Job time: 18199 sec

THIS PAGE BLANK (USPTO)